



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

Address: COMMISSIONER FOR PATENTS

P.O. Box 1450

Alexandria, Virginia 22313-1450

www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/526,810	03/04/2005	Nicholas James Midgley	GB920020049US1	8811
26502	7590	08/01/2008		
IBM CORPORATION IPLAW SHCB/40-3 1701 NORTH STREET ENDICOTT, NY 13760			EXAMINER TAHA, SHAQ	
			ART UNIT 2146	PAPER NUMBER
			MAIL DATE 08/01/2008	DELIVERY MODE PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary**Application No.**

10/526,810

Applicant(s)

MIDGLEY, NICHOLAS JAMES

Examiner

SHAQ TAHA

Art Unit

2146

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --
Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☒ Responsive to communication(s) filed on 21 May 2008.
2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 22 - 37 is/are pending in the application.
4a) Of the above claim(s) _____ is/are withdrawn from consideration.
5) ☐ Claim(s) _____ is/are allowed.
6) ☒ Claim(s) 22 - 37 is/are rejected.
7) ☐ Claim(s) _____ is/are objected to.
8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☐ The specification is objected to by the Examiner.
10) ☒ The drawing(s) filed on 04 March 2005 is/are: a) ☒ accepted or b) ☐ objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
a) ☒ All b) ☐ Some * c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
2. ☐ Certified copies of the priority documents have been received in Application No. _____.
3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- 1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO-8508)
Paper No(s)/Mail Date _____
4) ☐ Interview Summary (PTO-413)
Paper No(s)/Mail Date _____
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____

DETAILED ACTION

In view of the Appeal Brief filed on 05/21/08, PROSECUTION IS HEREBY REOPENED. New grounds of rejection are set forth below. Any inconvenience is regretted.

Claim Rejections - 35 USC § 103

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

Claims 22, 23, 26, and 36 are rejected under 35 U.S.C. 103(a) as being unpatentable over Noland et al. (US 7,080,378), in view of Bertram et al. (US 2003/0014507).

Regarding claims 22, 26, and 36, a method for allocating servers to a cluster of servers, said method comprising the steps of:

based on the performance data, said second server determining if said first server has reached a predetermined upper level of utilization, **[the controller software determines if the currently deployed virtual servers are approaching their defined service capacity, (Noland et al., Col. 5, Lines 5-8)],**

and if said first server has reached said predetermined upper level of utilization, **[if the deployed virtual servers are close to near-term service saturation, wherein**

the server has reached an upper utilization level, (Noland et al., Col. 5, Lines 27-28)],

said second server automatically sending a reconfiguration request to a server responsible for allocating servers to said cluster to allocate another server to said cluster, and in response, said responsible server automatically identifying another, available server and connection information for said other server and automatically allocating said other server to said cluster, **[the software will create and deploy a new virtual server with identical applications (step 504), and then direct the service request to this new server (step 505). This new virtual server can be deployed and activated, (Noland et al., Col. 5, Lines 28-35)],**

Noland et al. fails to teaches sending performance data regarding a server in a cluster,

Bertram et al. teaches providing performance analysis on a system including a cluster. The cluster includes a plurality of nodes. The method and system comprise obtaining data for the plurality of nodes and analyzing the data, **(Bertram et al., Paragraph 6, Page 1)**, to determine whether performance of the cluster can be improved, **(Bertram et al., Paragraph 6, Page 1)**,

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to modify Noland et al. by sending performance analysis on a system including a cluster. The cluster includes a plurality of nodes. The method and system comprise obtaining data for the plurality of nodes and analyzing the data,

(Bertram et al., Paragraph 6, Page 1), to determine whether performance of the cluster can be improved, (Bertram et al., Paragraph 6, Page 1).

Regarding claim 23, a method as set forth in claim 22 wherein the step of automatically allocating said other server to said cluster comprises the steps of updating a configuration file of said responsible server to list said other server as part of said cluster, **[the software will create and deploy a new virtual server with identical applications (step 504), and then direct the service request to this new server (step 505). This new virtual server can be deployed and activated, (Noland et al., Col. 5, Lines 28-35)]**.

Claims 24, 25, 27, 32, and 37 are rejected under 35 U.S.C. 103(a) as being unpatentable over Noland et al. (US 7,080,378), in view of Mashayekhi et al. (US 2003/0051187).

Regarding claims 24, 27, 32 and 37, Noland teaches the software will create and deploy a new virtual server with identical applications (step 504), and then direct the service request to this new server (step 505). This new virtual server can be deployed and activated, **(Noland et al., Col. 5, Lines 28-35)**,

Noland et al. fails to teach that if a server is under utilized then said server is no longer needed in said cluster,

Mashayekhi et al. teaches a method as set forth in claim 22 further comprising the steps of: based on the performance data, said second server determining if said first server is functional but under utilized such that said first server is no longer needed in said cluster, a determination is then made at step 252 as to whether node n+1 or node n-1 has the lowest weight, If node n+1 has the lowest weight, that node is established as the failover node (step 253), **(Mashayekhi et al., Paragraph 39, Page 4)**, to choose a more intelligent failover node, thereby improving the high availability of the cluster system, **(Mashayekhi et al., Paragraph 39, Page 4)**,

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to Noland et al. by determining if said first server is functional but under utilized such that said first server is no longer needed in said cluster, a determination is then made at step 252 as to whether node n+1 or node n-1 has the lowest weight, If node n+1 has the lowest weight, that node is established as the failover node (step 253), **(Mashayekhi et al., Paragraph 39, Page 4)**, to choose a more intelligent failover node, thereby improving the high availability of the cluster system, **(Mashayekhi et al., Paragraph 39, Page 4)**.

Regarding claim 25, a method as set forth in claim 24 wherein the step of automatically de-allocating said first server from said cluster comprises the step of updating a configuration file of said responsible server to remove said first server from said cluster, **[if a virtual server reaches a deactivation delimiter (e.g., it remains idle for the specified length of time), it is no longer needed to handle the current**

or anticipated workload. The control software removes it from the operating system environment (step 509), (Noland et al., Col. 5, Lines 51-53)).

Claims 28 – 31, and 33 are rejected under 35 U.S.C. 103(a) as being unpatentable over Noland et al. (US 7,080,378), in view of Bertram et al. (US 2003/0014507), and further in view of Dutta et al. (7,177,901).

Regarding claim 28, a method for managing servers, said method comprising the steps of:

a first server determining performance data for said first server and performance data for a second server, **[The Control Program 201 will dispatch virtual machines from an "eligible to run" list based upon various parameters, wherein server 201 determines the performance of other servers in the cluster, (Noland et al., Col. 3, Lines 33-35)],**

based on the reported performance data, said third server determining if said first server or said second server has reached a predetermined upper level of utilization, **[if some predefined resource in the virtual servers, such as CPU usage or a request queue into the cluster of servers, should reach a critical threshold, then one or more new virtual server(s) are automatically deployed, (Noland et al., Col. 1, Lines 65-67)];**

Noland et al. fails to teaches sending performance data regarding a server in a cluster,

Bertram et al. teaches providing performance analysis on a system including a cluster. The cluster includes a plurality of nodes. The method and system comprise obtaining data for the plurality of nodes and analyzing the data, **(Bertram et al., Paragraph 6, Page 1)**, to determine whether performance of the cluster can be improved, **(Bertram et al., Paragraph 6, Page 1)**,

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to modify Noland et al. by sending performance analysis on a system including a cluster. The cluster includes a plurality of nodes. The method and system comprise obtaining data for the plurality of nodes and analyzing the data, **(Bertram et al., Paragraph 6, Page 1)**, to determine whether performance of the cluster can be improved, **(Bertram et al., Paragraph 6, Page 1)**.

The modified Noland et al. fails to teaches reducing subsequent utilization of a server if it reaches a predetermined upper level of utilization,

Dutta et al. teaches a central load distribution server that is designed to redirect file requests to specific servers and protect any one server from an overload of traffic, **(Dutta et al., Col. 2, Lines 25-27)**, to allow a load distribution server to protect any given content server from being overload with file requests, **(Dutta et al., Col. 4, Lines 46-49)**,

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to modify the modified Noland et al. by reducing subsequent utilization of a server if it reaches a predetermined upper level of utilization, a central load distribution server that is designed to redirect file requests to specific servers and

Art Unit: 2146

protect any one server from an overload of traffic, **(Dutta et al., Col. 2, Lines 25-27)**, to allow a load distribution server to protect any given content server from being overload with file requests, **(Dutta et al., Col. 4, Lines 46-49)**.

Regarding claim 29, a method as set forth in claim 28 further comprising the earlier steps of:

said first server receiving a request from a client device and determining whether said first server should handle said request, and if so, said first server handling said request,

[When a new service request is received (step 501), the controller software determines if the currently deployed virtual servers are approaching their defined service capacity (step 502),, (Noland et al., Col. 5, Lines 5-7)],

and if not, said first server identifying said second server as available to handle said request and forwarding the request to said second server for handling, **[If the deployed virtual servers are not at the defined saturation threshold, the controller software directs the new service request to one of the deployed servers (step 503), (Noland et al., Col. 5, Lines 24-26)].**

Regarding claim 30, a method as set forth in claim 28 wherein the step of said first server automatically reconfiguring itself comprises the step of said first server updating a configuration file which lists one or more servers which are available to handle specified types of client requests, **[balancing workload among virtual servers configured in a virtual server cluster, wherein workload is balanced among**

different servers depending on its availability, (Noland et al., Col. 6, Lines 18-21)].

Regarding claim 31, a method as set forth in claim 28 wherein the step of said first server automatically reconfiguring itself comprises the step of automatically identifying a fourth, available server and connection information for said fourth server and automatically allocating said fourth server to said cluster, **[automatically deploying a new at least one additional virtual server to the virtual server cluster such that the at least one additional virtual server is added to and operates in tandem with the virtual server cluster, (Noland et al., Col. 6, Lines 26-29)]**.

Regarding claim 33, a method as set forth in claim 32 wherein the automatic de-allocating reconfiguring step comprises the step of adding to a pool of available servers, said server which is functional but under utilized and no longer needed in said cluster, **[the software will create and deploy a new virtual server with identical applications (step 504), and then direct the service request to this new server (step 505), (Noland et al., Col. 5, Lines 28-35)]**,

such that if a server in said cluster subsequently reaches said predetermined upper level of utilization, the de-allocated server can be re-allocated to said cluster, **This new virtual server can be deployed and activated, (Noland et al., Col. 5, Lines 28-35)]**.

Claims 34 and 35 are rejected under 35 U.S.C. 103(a) as being unpatentable over Noland et al. (US 7,080,378), in view of Bedell et al. (US 6,859,798).

Regarding claim 34, Noland teaches the software will create and deploy a new virtual server with identical applications (step 504), and then direct the service request to this new server (step 505). This new virtual server can be deployed and activated, **(Noland et al., Col. 5, Lines 28-35)**,

Noland et al. fails to teaches sending data from first server to a second server using XML data stream,

Bedell et al. teaches a user can send a request to the intelligence server 300 in an XML string, and the intelligence server 300 will send back its result in XML as well, **(Bedell et al., Col. 10, Lines 5-9)**, to provide a more open communication protocol, **(Bedell et al., Col. 10, Lines 3-4)**,

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to modify the modified Noland by sending data from first server to a second server using XML data stream, wherein a user can send a request to the intelligence server 300 in an XML string, and the intelligence server 300 will send back its result in XML as well, **(Bedell et al., Col. 10, Lines 5-9)**, to provide a more open communication protocol, **(Bedell et al., Col. 10, Lines 3-4)**.

Regarding claim 35, Noland teaches the software will create and deploy a new virtual server with identical applications (step 504), and then direct the service request to this new server (step 505). This new virtual server can be deployed and activated, **(Noland et al., Col. 5, Lines 28-35),**

Noland et al. fails to teaches determining if said request is for data which is currently cached at said first server,

Bedell et al. teaches cache all prior user-requested objects. In this architecture, most calls to a cached object will be handled within the user engine 102 to achieve the best performance, **(Bedell et al., Col. 11, Lines 19-22)**, to minimizes the client/server communication, **(Bedell et al., Col. 11, Lines 16-18),**

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to modify the modified Noland by determining if said request is for data which is currently cached at said first server, wherein cache all prior user-requested objects. In this architecture, most calls to a cached object will be handled within the user engine 102 to achieve the best performance, **(Bedell et al., Col. 11, Lines 19-22)**, to minimizes the client/server communication, **(Bedell et al., Col. 11, Lines 16-18).**

Conclusion

Any inquiry concerning this communication or earlier communications from the examiner should be directed to **Shaq Taha** whose telephone number is 571-270-1921. The examiner can normally be reached on 8:30am-5pm Mon-Fri.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, **Jeff Pwu** can be reached on 571-272-6798.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

/S. T./

Examiner, Art Unit 2146

/Jeffrey Pwu/

Supervisory Patent Examiner, Art Unit 2146